DOCUMENT RESUME

ED 126 867 IR 003 772

AUTHOR Mielke, Keith W.; Bryant, Jennings, Jr.
TITLE Formative Research in Attention and Appeal: A Series of Proposals.
INSTITUTION Children's Television Workshop, New York, N.Y.
PUB DATE 11 Jul 72
NOTE 21p.; Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Attention; Childhood Interests; Day Care Services; Educational Television; Formative Evaluation; *Interest Tests; *Measurement Techniques; Research Methodology; *Television Research; *Television Viewing
IDENTIFIERS Childrens Television Workshop; Distractor Technique

ABSTRACT

Methods are suggested to measure the program appeal and audience attention of Children's Television Workshop productions. Among these are distractor techniques, one which permits subjects to discriminate between two simultaneously broadcast programs by selecting the audio track they most prefer and one used to rank order several programs. Requiring the subject to push a button to maintain the broadcast and then comparing the button pushing frequency of different programs; using physiological measures of excitement or arousal; and allowing children in day care centers to voluntarily select television viewing amidst a multitude of controlled distractions are also suggested. The final suggestions include sequential photographs of subjects in natural settings, profile analysis, infra-red photography of subjects viewing television in a dark room, and binocular testing which places a different image in front of each of the subjects' eyes and records which image is most captivating. (EMH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED126867

FORMATIVE RESEARCH

IN ATTENTION AND APPEAL

A Series of Proposals

Keith W. Mielke
Jennings Bryant, Jr.
Children's Television Workshop
July 11, 1972

This document was processed for the ERIC Document Reproduction Service by the ERIC Clearinghouse at Stanford. We are aware that some pages probably will not be readable in microfiche ~~as in Hardcopy~~ form. However, this is the best available copy, and we feel that the document should not be withheld from interested readers on the basis of these unreadable pages alone.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HR003772

INTRODUCTION

In an earlier paper we presented some ideas for formative research in the area of comprehension.* In this paper, the related prerequisites of program appeal and audience attention are considered as a separate set of formative research problems. Parts of the paper contain suggestions for possible improvement of current CTW formative research practices in program appeal (e.g., modifying the distractor technique to allow testing in groups). Other designs are not, to our knowledge, in use at CTW. As with the earlier paper, these are presented as suggestions with the objective of stimulating discussion and constructive criticism, in the hope that at least some of the ideas will merit field testing and adoption by the CTW research staff.

I. A Group Testing Method for the Distractor Technique

The distractor technique has become a standard and useful procedure in CTW formative research, but the restriction to a single subject at a time means that data collection is expensive and time-consuming. As a consequence, sample sizes are often disturbingly small. The procedure suggested here, if it works, would generate comparable distractor data with tests groups of about twenty subjects.

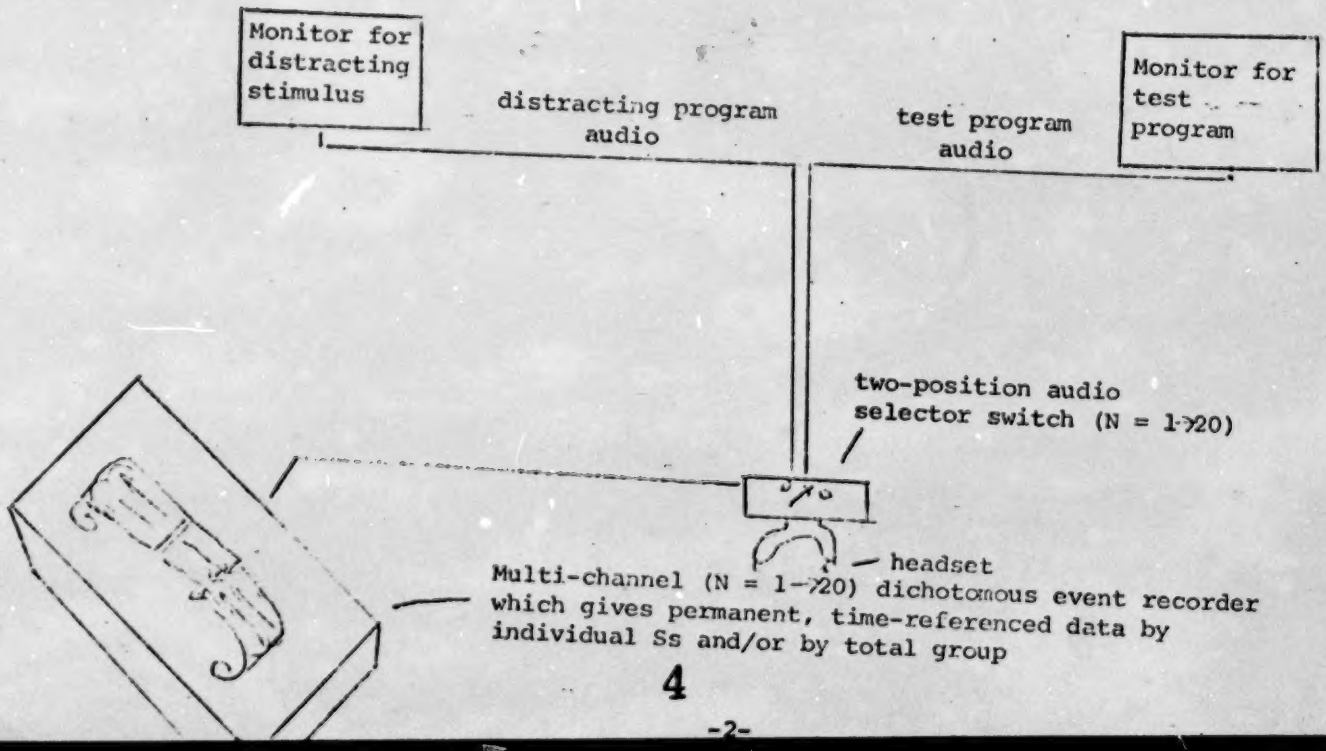
---

*Keith W. Mielke and Jennings Bryant, Jr., "Formative Research in Comprehension of CTW Programs: A Series of Proposals," Children's Television Workshop, June 30, 1972.

Data that now take a week or more to collect could be gathered in one day, making "fast reaction" service to production a reality. Expenses should decrease and sample sizes should increase. Distractor data now hand tabulated from coder sheets would be recorded automatically, and could be displayed in the familiar "attention profile" format. As with the present distractor technique, the modified version suggested here can become a standardized procedure into which any new program or program segment can be incorporated.

The critical modification from individual to group testing lies in what specific behavior is recorded as data. The current method uses sequential observer notations of a single subject's direction of looking (toward the test program or toward the distracting stimulus). The proposed method uses a distracting stimulus that has an audio track; the measured behavior is the sequential selection of audio (from the test program or from the distracting stimulus).

A schematic diagram of the apparatus needed is given below:


Monitor for distracting stimulus
distracting program audio
test program audio
Monitor for test program
two-position audio selector switch (N = 1->20)
headset
Multi-channel (N = 1->20) dichotomous event recorder which gives permanent, time-referenced data by individual Ss and/or by total group

The use of television "programs" with audio and visual content as distracting stimulus material, poses some problems not encountered in the use of still slides as distractor:

A. The dynamic distracting stimulus could have varying degrees of attractiveness internally, i.e., the reference point is not inherently stable.

B. Initial subject involvement in a "distracting" story line could hold his attention on the distractor even through uninteresting portions. That is, a subject might "commit" to one monitor or the other and thereafter not even consider the alternative stimulus.

Proposed solutions to these problems will be discussed next.

What kind of distracting stimulus material should be selected? Given the various lengths of CTW test material, it seems that short (30 to 90 second) commercials would be ideal distractor material. Combinations of commercials could be made to any length. If, for some reason, the selection of any commercial were a bad choice, the negative or invalidating effects of this error would not be of long duration (e.g., Ss could not get "hooked" for long periods of time). A large pool of distracting stimuli (commercials) should be readily available in almost any production category.

How could we determine and measure how distracting the distracting stimulus is? What is the attention-getting power of the distracting stimulus? Ultimately, this question is equally valid for the still slides now used as distractors, but the issue does not arise there because the

slides are considered (perhaps erroneously) a "constant." A "constant" can be a useful reference point for relative measurements, even if the location of the constant is unknown. The known variability of commercials, however, makes this an issue that must be grappled with. The suggestion is to process a large number of commercials through the traditional distractor methodology, obtaining quantitative attention profiles on each. A set of commercials, matched on the traditional attention indices, would form the population of distractor stimuli for the proposed methodology. Intuitively, the preference would be for commercials matched at a low level of appeal for children.

What methodological features can be devised to assure that the subject does not "commit" to one monitor and simply ignore the other? Again, the issue does not arise with the traditional procedure, because the sequence of slides is not psychologically connective. The slides are unlikely to have the power of "hooking and holding" sustained attention. The brevity of the segments within the distracting stimulus in the proposed methodology reduces the danger somewhat, but additional controls are possible. The suggestion is to switch or alternate the signals between the monitors at reasonable time intervals. A reasonable interval would be the shortest interval that does not produce a dysfunctional degree of frustration or irritation among the subjects. The interval would have to be determined empirically. Every "N" minutes, the picture and sound on the left monitor would switch to the right monitor, and vice versa. In a "perfect" CTW program, all Ss would switch their audio selector to continue attending the CTW program. In other words, a conscious decision-making situation is imposed on the Ss every "N" minutes. A "non-coping" S (one who left the audio switch unmanipulated

throughout) could simply be removed from the analysis.

Once this procedure is routinized, it could be set up to process twenty subjects with approximately the same effort now required to process one subject.

An innovative engineer could automate all of the signal switching and connect this with the event recorder so that data transformation would be unnecessary.

Variations of this technique may also be of interest, such as pitting one CTW segment against another, or testing two production versions of the same segment. One variation might be measuring the strength of appeal of the bit by the length of "interest carryover effect" that it has. In this design Monitor A (whichever monitor carries CTW material at the conclusion of the bit) would then carry inserted boring material of known appeal strength immediately after the test bit was finished. Monitor B would continue carrying the distracting material. The dependent measure (the interest carryover) would be operationalized as the amount of time the S continued to "prefer" Monitor A after the boring material was introduced. Another variation utilizing this apparatus is described in Design II.

II. Paired Comparison Measures Utilizing Viewing Behavior Instead of Paper-Pencil Tests

Traditionally, the paired comparisons technique requests systematic preference choices between two alternatives throughout all combinations of alternatives. With a pool of four items, preference

selections would be made within (N) (N-1) /2 or six pairs. The index of subject preference in this design would be actual viewing behavior as recorded with the apparatus described above in Design I. "Items" would be program segments. Preference choices would be between simultaneously presented pairs of segments. The data would indicate the rank order of preferability among the pool of items through tabulation of the several dichotomous choices.

To illustrate, consider the case of four program segments, A, B, C, and D, for which a rank order of preferability is desired. Presentations could be made as follows:

	Left Monitor	Right Monitor
1.	A	B
2.	C	D
3.	A	C
4.	B	D
5.	D	A
6.	B	C

For each subject, the preferred member of each pair could be determined simply by which alternative received the majority of his attention (as indexed by audio selection). Preferences would be tallied as follows, where dummy data for twenty Ss are given:

	More Preferred			
Less Preferred	A	B	C	D
A	XXX	8	15	11
B	12	XXX	18	14
C	5	2	XXX	7
D	9	6	13	XXX
TOTALS:	26	16	46	32

In this hypothetical example, the most preferred segment is "C," followed by "D," "A," and finally "B."

Certain methodological problems in Design II should be considered. For one thing, the segments ideally should be of exactly the same length, and this is not likely without extensive editing. The suggestion would be to let all segments run their entire length, but to restrict data analysis to the length of the shortest segment in the group. Also, the repeated use of program segments in the various pairings brings up extraneous variables such as satiation. This problem can be met in two ways. One would be to ignore the repetition effects, using the argument that all segments are repeated the same number of times; hence, any depressant bias effect would be uniform. Repeated use of the same bits would thus yield bit-specific data, in contrast to an alternate method of using different exemplars for a bit by type:

A (version 1)	B (version 1)
A (version 2	C (version 1)
B (version 2)	C (version 2)
Etc.	

This latter method would, of course, yield rank order data on bit types rather than on specific bits. Intuitively, the former method (repeated use of same bits) seems preferable.

It may be possible to get more precise measurement than "greater than/less than" relationships from the apparatus of Design I used in the paired comparisons of Design II. Assume that the continuous data from the apparatus would be "read" every two seconds, or thirty times per minute. Assume comparative testing of four one-minute bits,

using twenty subjects. As seen in the earlier diagram, each bit would be played three times, so the maximum score for a bit in this situation would be 30 (measures per minute) X 1 (minutes per bit) X 3 (repetitions) X 20 (subjects) = 1800. Each bit could therefore be scored on the percentage of 1800 actually achieved, thus yielding not only a rank order, but also preferential distance points between the bits.

III. Behavioral Response Required to Activate/Maintain Audio or Video Channel

The key to this design is a device which requires effort on the part of the S to maintain an intelligible stimulus level. The amount of effort the S will exert to maintain the stimulus is the correlate measure of the appeal of the bit.

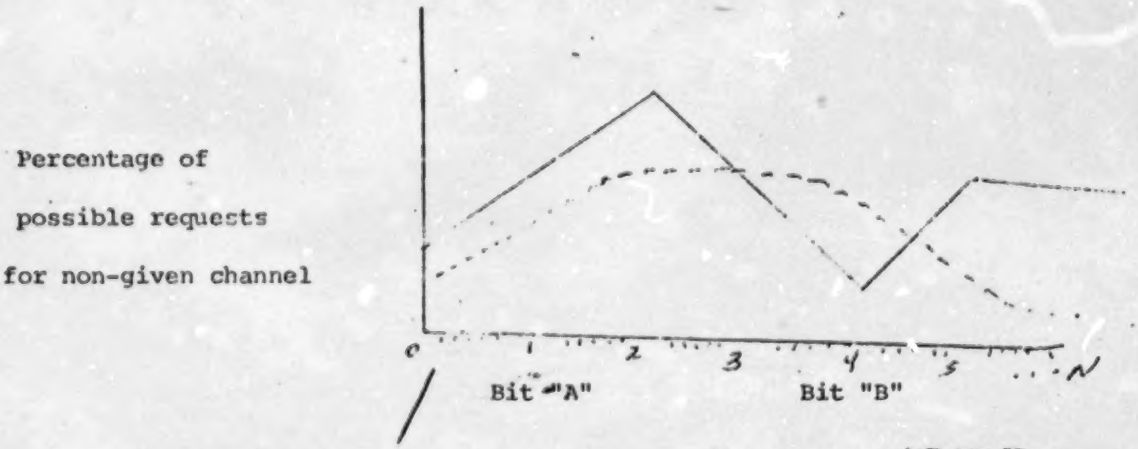
Procedurally, the testing could go as follows:

1. S is seated in normal television viewing position.
2. Videotaped stimulus material (CTW bit to be evaluated) is presented.
3. In order to maintain audio channel (or video channel), S must press a button at an established interval (every ½ to every 2 seconds). A cued time-line polygraph recorder indicates the number of button presses and the exact time of the response.

If the effectiveness of the audio-deletion method proves to be equal to the video-deletion method, then Ss could be tested in groups, thus providing great economy. Earphones and multiple polygraph recorders would be the only additional equipment requirements, and no additional personnel would be needed.

To combat the anticipated novelty effect of pushing the button regardless of what is on the screen, a "novelty effect extinction period" is proposed, wherein boring material would be shown until the button is no longer pushed. Then and only then would the testing begin. As a validity check on subsequent button-pushing behavior, we would recommend periodic insertion of material known to be boring. A model subject viewing a model CTW program would stop pushing the button shortly after the boring insert began and would resume pushing shortly after the CTW program resumed. Field data would be needed to establish norms for extinguishing novelty effect behavior and for establishing the criteria for a non-coping subject, to be deleted from the analysis.

The data presentation envisioned could be as follows:



_____ = requests for audio, given video

_ _ _ _ _ _ _ = requests for video, given audio

The table above contains dummy data on both audio and video requests, which may be an unnecessary luxury if the request curves correlate highly. If request patterns are strikingly different, however, both sets of data should be gathered and interpreted. Presumably all

sets of audio-video request data could be placed in one of the following cells:

Audio Requests, Given Video

Video Requests, Given Audio		HIGH	LOW
	HIGH	CELL A	CELL B
	LOW	CELL C	CELL D

Cell A locates unambiguously a high appeal program; Cell D locates unambiguously a low appeal program; it is in cells B and C that inconsistencies require more elaborate interpretation. The particular strength of this design lies in its ability to provide channel-specific behavioral data. This capability, when tied in with channel-specific content analysis data, would appear to make this technique a powerful diagnostic tool for inconsistencies found in Cells B and C.

Explanation of audio-video inconsistencies is a "bonus spin-off;" the main use for which Design III is intended is a direct test of program appeal as indexed by the amount of effort a subject is willing to exert to receive it.

IV. Physiological Measures

Physiological measures offer several advantages:

A. They are nonverbal, and, in some cases, involuntary.

B. They are generally reliable (but their validity frequently

requires cross-referencing with other measures).

C. Depending on the outlay for equipment, Ss can be tested individually or in groups.

D. In recent years equipment costs have declined and reliability has increased.

E. Physiological measures can go beyond attention to assess some aspects of stimulus impact.

In particular, physiological measures should be considered as CTW moves into the domain of affective program objectives, where verbal measures are more difficult (and possibly less relevant) than they are with current programming.

Physiological measures of arousal include Galvanic skin response (GSR), heartbeat, and blood pressure readings. As the subject views the CTW test material, continuous time-referenced physiological measures would allow exact synchronization with stimulus material.

Arousal measures tend to be non-discriminative; i.e., we can measure arousal, but we can't determine if the arousal occurred because of increased bit appeal or because of annoyance with the bit, or because of some extraneous source. It is suggested, therefore, that observational measures be taken concurrently with physiological measures as an interpretive aid.

The degree to which the apparatus and the "wiring in" will produce subject apprehension must be empirically determined.

Another physiological measure that has the advantage of being a completely unobtrusive measure of bit appeal relies on a sensitive

device known as a "wiggle meter." The wiggle meter measures the Ss' amount of movement as they observe stimulus material, a measure which may be particularly germane to our youthful target audience. Measurement, recorded on a time-line polygraph recorder, is continuous and sensitive.

The procedure is simple and can be adapted to almost any viewing situation.

A. Ss are seated in special "wired" chairs (wires are concealed in the seats of the chairs and cannot be felt by the subjects) Testing is performed in otherwise normal viewing conditions. Special care is taken to provide a normal, relaxed, nonexperimental atmosphere. No equipment is visible.

B. A baseline reading is obtained by having Ss view "neutral" stimulus materials.

C. Ss view stimulus material to be tested. Constant readings are taken throughout the viewing.

D. Post-tests are made, once again utilizing neutral stimulus material, to index the amount of normal fatigue generated during the testing period.

An intuitive assumption made in utilizing this design is that S movement is an index of the appeal strength of the bit. Children watching TV are generally quite wiggly observers; however, they generally wiggle (squirm) more when they are bored than when they are interested. This assumption is the basis for the expectation that an increase in wiggliness of a S is inversely proportional to the amount of appeal of the bit.

This hypothesis is valid only within limits. Ss may jump up and down with excitement when they are really "turned on" by a bit; however, this can be recorded in a decodable manner by the wiggle meter.

Individual differences of viewers must be considered; therefore, baseline data on each S must be obtained. Luckily, this is a rather easy task.

S fatigue is another problem. A high-appeal bit which occurs early in a presentation will probably result in less wiggling response than an equally appealing bit appearing later in the presentation. We can account for this by using short testing periods and by running post-tests which measure interest extinction due to fatigue.

Advantages in this procedure are that it can give the producer appeal data which was obtained unobtrusively from a large number of Ss viewing the stimulus material simultaneously in a "normal" viewing situation.

It is realized that formative research is not designed to test hypotheses, and the suggestions above are primarily for the benefit of producers. Some of the theoretic potential of these measures should at least be mentioned, however. Of particular interest might be the amount of physiological arousal that is most inducive to cognitive achievement. The decay curves of arousal could thus influence the placement of arousing bits within the program. "Filler" material (segments inserted primarily for their entertainment value) may function as arousal agents that facilitate subsequent learning at certain points in the decay curve. Demonstration of this effect and establishment of norms for it would ultimately have quite practical implications (if demonstrated, it could also serve as a data-based reply to those who criticize the entertainment content in CTW programs).

V. Voluntary Exposure to CTW Programs Amidst Competing Attractions in Day Care Centers

Going directly to target audiences in day care centers, this design

features portable, individualized, soundproofed viewing booths. They would be equipped with small monitors connected with video cassette playback units. CTW test programming would be played continuously during operating hours of the school. A time-referenced event recorder would automatically keep track of which program(s) were running at what times. Pressure-sensitive seats in the viewing booths would also record automatically when viewing took place.

An observer would keep a continuous content analysis of what the competing activities were, with time references. Cross tabulation of exposure data with competing activities data would, over time, give appeal indices under natural conditions for the school-based audience.

As more viewing booths and playback units are added, the design becomes increasingly flexible and powerful. Programs from the competition could be played to see how well CTW programs stand up under that realistic situation. One CTW program could be pitted against another for purposes of comparison. A set of programs could be continuously repeated, without variation, to see how long it takes for total extinction of viewing motivation to take place. The observer could note which non-viewing activities seem to reflect the influence of CTW programming.

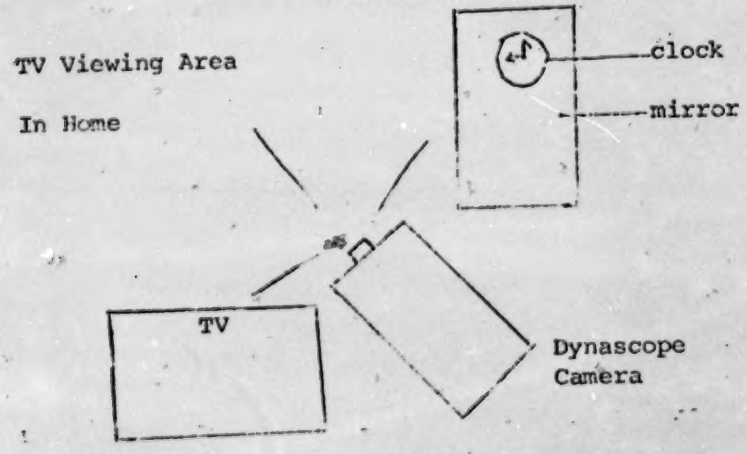
The advantages of this design are that it permits longitudinal testing in a non-test environment (no test anxiety), with no verbal measures involved, but with measures of actual viewing recorded precisely.

With a little imagination on our part, and cooperation from the classroom teacher, it should be possible to develop additional sophistication in the natural testing situation. For example: a certain amount of play money could be given to each child at the beginning of the day. Each activity in the school would "cost" the child a certain amount.

The "price" of the opportunity to view CTW programming could be manipulated experimentally, giving a gross quantitative measure of CTW program appeal, relative to other attractions in the child's school environment.

VI. Analysis of Sequential Photos of the Audience in Natural Settings

This is one of the two designs that permit testing in the natural home environment. (The first was reported in our paper of June 30, 1972: Design VII, "Audio Monitoring of Children's Audio Behavior Throughout the Day."). The idea suggested here was reported in the Journal of Advertising Research several years ago, where the apparatus, under the brand name of "Dynascope," was described as follows:



The Dynascope takes sequential still photos at a variable and controllable rate. Wide angle lenses take in not only the viewing audience, but also a mirror and clock, which thus gives a photographic record of the TV image and the exact time as well. Whenever the TV is turned on, the camera is activated.

If ethical arrangements can be made for using such a device in homes, a wealth of data will become available, answering such questions

as:

1. What are the ages of the actual CTW audience?
2. What size are the natural viewing groups?
3. What other activities coincide with viewing?
4. What is the total TV diet for the family?
5. Which CTW segments hold attention and which do not?
6. How active or passive do the viewers appear?
7. How good is the reception?

This list could be expanded considerably. The generic question is: What can one learn about CTW programming through unobtrusive observation in the home, with full knowledge and consent of the persons observed, that yields a permanent, sequential visual record of the viewers and the program? In addition to complete explanation of the device and the use of the data, we recommend attractive payments to host homes for their cooperation.. Data gathered in this design are expensive, but are not replicable in any known laboratory situation.

VII. Program Analyzer Technique (Profile Analysis)

This hoary research technique is quite simple in concept: it consists of a continuous "question" asked of the subjects throughout a program, "answers" to which are given continuously by means of a push-button device. Typically, each subject has two buttons, one for each hand; these buttons permit a two-option response over time. A time-referenced event recorder yields a permanent record of all responses, which are then matched with the accompanying stimulus material.

The data are obviously specific to the question asked, and a considerable variety of questions is feasible, at least with adult

audiences. For example:

A. Are you interested in the program? (push yes or no button)

B. Do you understand what is going on?

C. Are you learning?

With children, it is intuitively apparent that such self-diagnostic questions as "Are you learning?" would not be appropriate. It may be possible, however, to ask some simple questions in the affective domain and get useful data. For example, a child might be able to cope on a continuing basis with a question such as, "Are you happy now?" The push buttons could be presented as smiling and frowning faces. If this would work, it would be a very inexpensive means of collecting affective data.

VIII. Infrared Photography of the Audience in Group Settings

It is, of course, easy and inexpensive to use a regular TV camera for monitoring TV viewing behavior in settings such as day care centers. After a short period of time, children become oblivious to the camera and behave "normally." The semi-daylight conditions of such viewing, however, still permit group interaction and may inhibit facial expression of affective states because of distractions. The main feature of this design (borrowed from film testing where light-levels must be low) is, for our purposes, not the mere feasibility of audience observation, because simple TV monitoring is easier. Instead, the rationale here is based on the different behavior displays presumably induced by darkened viewing conditions: less distraction, more attention to the screen, less self-conscious of physical appearance, more candor in facial expressions, less consciousness of the presence of others, and less opportunity for alternative activities. In darkness, the viewer is

"hooked" to the screen to a degree unmatched in normal TV viewing.

Infrared photography of the audience under conditions of darkness would permit such things as aiming the camera at a single subject and zooming in for a closeup with no awareness or reaction on the part of the subject. If infrared TV cameras are available, they would be ideal because of their economy, and the fact that audio pickup of the program's soundtrack would allow precise matching with the stimulus materials.

Instead of attention measures, which could be collected more economically through other means, this design seems particularly appropriate for attempts at measurement of affective states as displayed facially, capitalizing on the effects of darkness.

IX. Binocular Testing

One method of assessing what a person is physchologically prepared to see is by means of binocular testing. Such a question could well be relevant to CTW producers as they get increasingly involved in cultural minority programming. Any time that the effects on perception of enculturation presents a suspected problem in production decisions, binocular testing or some variation might be useful.

A physical divider, extending from the nose to the stimulus, separates two images so that each eye sees a different stimulus. The subject is asked to describe what he sees. Subject response to presentation of production alternatives would indicate the version most readily perceived. Subject response could also be evaluated as in a quasi-projective test, i.e., to permit inferences on the subject's connotative

meaning for the stimulus. Effects of left- or right-eye dominance could be controlled in several ways, such as alternation across two exposures.